DOCUMENT RESUME

ED 146 038 SE 023 240

AUTHOR	Blake, Anthony J. D.
TITLE	Measurement of Level of Intellectual Development in Australian Secondary Schools: A Report to the Education Research and Development Committee.
INSTITUTION	Riverina Coll. of Advanced Education, Wagga Wagga, New South Wales (Australia).
SPONS AGENCY	Australian Education Research and Development Committee, Canberra.
PUB DATE	May 77
NOTE	64p.
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	Achievement; *Cognitive Development; *Educational Research; Intellectual Development; Learning; *Measurement Instruments; *Reading Comprehension; Science Education; Secondary Education; *Secondary School Science
IDENTIFIERS	*Australia

ABSTRACT

This document reports data on 610 secondary level Australian high school students drawn from 36 urban and rural high schools to determine levels of Piagetian cognitive development, provide a profile of field independence (FDI) across the urban population, validate the Understanding in Science Test instrument, and determine influence of reading comprehension and FDI on the validity of group tests. Results for both rural and urban populations indicated a continual growth in cognitive development over the secondary grades with approximately 70% of grade 12 students being measured at the early or complete formal levels. The Understanding in Science Test instrument was not a good predictor of cognitive level of development; reading comprehension and FDI were found to be significant factors influencing the results of group test results. (SL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED146038

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

A. J. D. Blake

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Riverina College
of Advanced Education

Measurement of level of intellectual development in Australian Secondary Schools

A Report to the Education Research and Development Committee

SE 023 240

Anthony J. D. BLAKE,
School of Teacher Education.

May, 1977.

2

MEASUREMENT OF LEVEL OF INTELLECTUAL DEVELOPMENT IN AUSTRALIAN SECONDARY SCHOOLS

A Report to the Education Research and Development Committee

Anthony J.D. Blake
School of Teacher Education
Riverina College of Advanced Education

May, 1977

ACKNOWLEDGMENTS

This study was initiated with the generous support of the Director-General of Education in New South Wales, the Director of Education, Riverina Region, the Director of Secondary Education in Victoria, and the Australian Academy of Science School Biology Project.

Doug Hill was heavily involved in the planning of the project and assisted with the testing. David Morgan and John Baird organized and helped conduct the Melbourne phase of the survey. Jane Louis provided a great deal of assistance with computer programming and Ken Ross assisted with sampling procedures.

The study could not have proceeded without the willing co-operation of the Principals, staff and students of the participating schools and the outstanding assistance provided by the interviewers listed in Appendix 3 who expended much of their own time in conducting the testing.

The project was funded by a grant from the Education Research and Development Committee.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
PURPOSE	3
PROCEDURES	
The Populations Under Study	4
Sampling Procedures	5
Measuring the Variables	7
Estimating Validity of the Group Tests	15
Determining Significance of Correlation Coefficients	16
RESULTS AND DISCUSSION	
Level of Intellectual Development in the Two Populations	19
Field Dependence-Independence in the Melbourne Population	24
Relationship Between FDI and Level of Intellectual Development	27
Validity of the Understanding in Science Test	27
Validity of the Sheehan Test	37
CONCLUSIONS AND IMPLICATIONS	
Level of Intellectual Development in the Australian Secondary School Population	43
The Applicability of the Understanding in Science Test and the Sheehan Test	46
REFERENCES	48
APPENDICES	
Appendix 1. Schools included in the Riverina sample and details of sample structure	53
Appendix 2. Schools included in the Melbourne sample and details of sample structure	54
Appendix 3. List of trained interviewers employed in administration of the Piagetian tasks	55

TABLES

			Page
Table	1.	Riverina sample structure by sex and year.	6
Table	2.	Melbourne sample structure by sex and year.	8
Table	3.	Scheme for scoring the Piagetian tasks.	10
Table	4.	Scale of scores on battery of tasks used for determining predominant level of intellectual operations.	11
Table	5.	Performance on individual Piagetian tasks and on total task battery.	17
Table	6.	Mean scores on the battery of Piagetian tasks for each school year in both populations.	18
Table	7.	Percentages of subjects at each Piagetian level for each school year of both populations with 95% confidence limits.	21
Table	8.	Point biserial correlation coefficients between sex and performance on the Piagetian task battery.	23
Table	9.	Mean scores on the Group Embedded Figures Test across school years.	26
Table	10.	Matrix of product-moment correlation coefficients between the Understanding in Science test and the Piagetian tasks for the Riverina population.	28
Table	11.	Matrix of product-moment correlation coefficients between the Understanding in Science test and the Piagetian tasks for the Melbourne population.	29
Table	12.	Chi-square values and contingency coefficients for tests of independence between categorisations on the Piagetian task battery and Understanding in Science test.	31

Page

Table 13. Means of reading comprehension scores for subjects in each cell of the contingency table of task categorisations against Understanding in Science test categorisations. 34

Table 14. Results of tests of significance of differences between means of reading comprehension scores within rows of the tasks categorisation, Understanding in Science test categorisation contingency table. 36

Table 15. Chi-square values and contingency coefficients for tests of independence between categorisations on the Piagetian task battery and on the Sheehan test. 38

Table 16. Means of reading comprehension scores for subjects in each cell of the contingency table of task categorisations against Sheehan test categorisations. 41

Table 17. Results of tests of significance of differences between means of reading comprehension scores within rows of the tasks categorisation, Sheehan test categorisation contingency table. 42

FIGURES

		Page
Figure 1.	Profiles of levels of intellectual development across school years for both populations.	20
Figure 2.	Frequency distributions of GEFT scores across school years.	25

INTRODUCTION

Over the past decade there has been a growing concern for the implications of Piaget's conception of intellectual development for curriculum construction and implementation. This interest has been particularly intense in science education. In Britain, the development of the Schools Council's 5/13 Project (Schools Council 5/13 Project, 1972) was overtly influenced by the Piagetian paradigm and there have been a series of investigations of the relationship between secondary science curriculum materials and Piaget's developmental stages (Ingle and Shayer, 1971; Shayer, 1970, 1972). The Science Curriculum Improvement Study materials produced in the United States are based on Piaget's conceptions of intellectual growth (Karplus and Thier, 1967) and in Australia, the Australian Science Education Project adopted the Piagetian description of intellectual development as a basis for the development of its materials (ASEP, 1974).

There are two fundamental features of a curriculum development based on the Piagetian model of intellectual growth. Firstly, there is a concern to provide each learner with a set of stimulating and challenging experiences which are appropriate for his current stage of development. Secondly, there is an intent to foster further development of each learner. It is necessary, therefore, that the teacher implementing such materials be able to assess the stage of intellectual development of each pupil. The method of clinical interview adopted by Piaget and his co-workers is a time-consuming procedure which is impractical for the teacher, so a need exists for a convenient group test. As Rowell and

Hoffmann (1975) stated, "What is required is a group test, easily administered, readily marked, and yet retaining as many as possible of the attributes of the original Piagetian tasks."

There are now a number of examples of attempts to produce such group tests. One of the earliest was a series of tests developed by Longeot (1962, 1965). A compilation of a series of items from these tests was prepared, translated and used by Sheehan (1970). This test is of particular interest because its content is of a general nature, not directly related to science, as are many of the Piagetian tasks themselves. Lawson and Blake (1976) found this so-called Sheehan test to be a valid instrument when used with American high school students, but Blake (1976) found that it was not so with a study of Australian college students. Clearly, this test needs further scrutiny and its validity is examined further in this study.

Karplus and Karplus (1970) and Karplus and Peterson (1970) described an "Islands Puzzle" and a "Ratio Puzzle" respectively, both of which were purported to assess stage of intellectual development. However, the validity of the Islands Puzzle as a measure of Piagetian stages has been questioned as the result of a study by Blake et al. (1976). More recent attempts to produce satisfactory group tests include those reported by Shayer and Wharry (1974) and Shayer et al. (1976) in Britain, and by Rowell and Hoffmann (1975) and Tisher and Dale (1975) in Australia. The Tisher and Dale test has reached the stage of being marketed under the title "Understanding in Science Test", although its validity has not been extensively investigated.

Piaget's theory of intellectual development deals with the

structure of the intellect and its growth. The logic implicit in the Piagetian tasks designed to diagnose stages of development is derived from this structural model. Yet a subject may be competent with the logic implicit in a task and not perform satisfactorily on it. Pascual-Leone (1969) has argued that performance on Piagetian tasks is influenced by factors other than stage of intellectual development. In particular, performance on some tasks is heavily influenced by the subject's level of field independence. Field dependence - independence (FDI) is a dimension of cognitive style described by Witkin et al. (1962). In essence, field independent subjects are less prone to be distracted or fazed by irrelevant stimuli in the field when they attend to a problem than are field dependent subjects. The influence of FDI on performance with cognitive tasks is supported by the studies of Case and Globerson (1974) and Niemark (1975). Pascual-Leone (1969) indicated that some Piagetian tasks, because of their nature, are more difficult for field dependent subjects than others. Consequently, in any investigation of performance on tests of level of intellectual development, a measure of level of FDI of the population and its influence on test performance, seems necessary.

PURPOSE

Although the Australian Science Education Project (ASEP) produced a series of curriculum materials for Australian schools which were designed to cater for differences in stages of intellectual development, it did so with very little

information about stages of development in the Australian secondary school population (ASEP, 1974). This lack of information is a significant disadvantage for curriculum development teams in all areas, not only science.

This study was designed to survey level of intellectual development across the secondary school years in two populations; an urban population and a rural population. This was seen as a primary objective, but in addition the following investigations were made.

1. In view of the apparent relationship between performance on problem-solving tasks and level of field independence, the study was designed to provide a profile of FDI across the secondary school years of an urban population.

2. The validities of the Understanding in Science Test (Tisher and Dale, 1975) and the Sheehan test were examined.

3. As both of the group tests involve subjects in reading, and given that the construct FDI potentially could influence performance on these group tests, the study was designed to determine whether or not these factors, reading comprehension and FDI, influence the validity of the group tests.

PROCEDURES

The Populations Under Study

Two populations were studied in this investigation, a rural secondary school population and an urban secondary school population. The rural population was that defined as the Riverina region in New South Wales and the urban

population was based on metropolitan Melbourne.

Sampling Procedures

1. The Riverina Population. A stratified random sample of 310 Ss was drawn from 16 government high schools, one central school and four independent schools. The final selection of schools was influenced by relative proximity to Wagga Wagga and the willingness of the particular school to take part in the study. However, the greatest proportion of the secondary population in the Riverina is contained within the 24 government high schools in the region. As 16 of these schools were included in the sampling, it is contended that the final sample represented a reliable indicator of the total population.

Within each school a sample was drawn of size proportional to the size of the school's enrolment in relation to the total population size. This sample was stratified on two variables, sex and year of school, and then subjects were drawn at random from class lists. The sample structure for each school is shown in Appendix 1, and the total sample breakdown by sex and year is shown in Table 1.

2. The Melbourne Population. A random sample of 300 Ss was drawn using a different procedure from that employed for the Riverina sampling. A two-stage stratified random sampling was performed, sampling first by schools and then by students within schools. These procedures followed those used in the IEA Science Project (Rosier and Williams, 1973) and the Literacy and Numeracy Study in Australia (Keeves and Bourke, 1976). Using a sampling frame developed by the Australian Council for Educational Research, schools were sampled with a

Table 1. Riverina sample structure by sex and year.

Sex	SCHOOL YEAR 7	8	9	10	11	12	TOTAL
M	29	30	32	27	29	16	163
F	26	27	28	24	27	15	147
TOTAL	55	57	60	51	56	31	310

probability proportional to their enrolment size. Fifteen schools were thus chosen and within these schools twenty students were chosen at random. As with the Riverina sampling, this was achieved by numbering students consecutively on class lists and then choosing the requisite number from tables of random numbers.

The sample structure for each school in the sample is shown in Appendix 2 and the total sample breakdown by sex and year is shown in Table 2.

3. Bias in the Sampling. An unavoidable element of bias was introduced into the Riverina sampling, although this did not occur in the Melbourne sampling. In a number of the Riverina schools, Principals would not allow the inclusion of Year 12 students in the study because of a perceived encroachment upon their preparation for external examinations. In such cases extra Year 11 students were substituted but this clearly affected the structure of the sample.

Measuring the Variables

1. Level of Intellectual Development.

(i) The Criterion Measure - A Battery of Piagetian Tasks. The criterion measure of level of intellectual development was a battery of Piagetian tasks composed of the following individual tasks:

(a) conservation of weight (Piaget and Inhelder, 1941);

(b) conservation of displacement volume (as interpreted by Karplus and Lavatelli, 1969);

(c) separation of variables (Inhelder and Piaget,

Table 2. Melbourne sample structure by sex and year.

Sex	SCHOOL YEAR 7	8	9	10	11	12	TOTAL
M	27	31	27	31	18	9	143
F	25	33	26	32	23	15	154
TOTAL	52	64	53	63	41	24	297

1958);

(d) equilibrium in the balance (Inhelder and Piaget, 1958).

This battery of tasks was administered to each S individually by a trained instructor following a set of standardised protocols. The details of these protocols have been given by Lawson (1973) and Blake (1976).

Scoring the Tasks. This study followed the practice adopted by similar investigations (e.g. Lawson, 1973; Lawson et al., 1975; Lawson and Blake, 1976; Blake et al., 1976) in employing a battery of tasks to measure level of intellectual development, rather than a single task. A battery provides a greater opportunity for a S to demonstrate competence. The four tasks employed were chosen for their relative speed of administration and clarity of protocol.

Subjects were scored separately on each task and responses categorised according to the scale in Table 3. Task 1, conservation of weight, measured to the Piagetian stage IIb (late concrete operations), so the maximum score on the task was three. Task 2, conservation of displacement volume, measured to stage IIIa (early formal operations), so the maximum score was four. The remaining two tasks measured up to stage IIIb (complete formal operations), so the maximum score on each was five. Hence, the total maximum score on the battery was 17.

The total score was taken as an indication of a S's tendency to operate at a particular Piagetian level and, for the purpose of characterising the populations, the scale in Table 4 was used to classify each S into one of the four Piagetian levels.

Table 3. Scheme for scoring the Piagetian tasks.

Score	Intellectual level	
1	I	preoperational
2	IIa	early concrete operations
3	IIb	complete concrete operations
4	IIIa	early formal operations
5	IIIb	complete formal operations

Table 4. Scale of scores on battery of tasks used for determining predominant level of intellectual operations.

Score on battery of tasks	Predominant level of intellectual operations
16 - 17	IIIb
14 - 15	IIIa
11 - 13	IIb
8 - 10	IIa
0 - 7	I

Inter-interviewer Consistency. In order to have all Ss in both samples tested at about the same time it was necessary to employ a relatively large number of interviewers. This being so, the following procedures were adopted in order to control inter-interviewer consistency.

Volunteers were solicited from among teacher education students at Riverina College of Advanced Education. Twenty-eight potential interviewers were instructed in the administration and scoring of the battery of tasks. Following this instruction, the interviewers were asked to score performances on the tasks of a series of Ss whose interviews had been recorded on video-tape. The scoring obtained by the author was used as the criterion and those interviewers whose scoring matched this exactly were retained. It is worth noting that, following initial training, inter-interviewer consistency was very high and 18 interviewers were retained. These students were used for testing in both populations, together with assistance from trained colleagues and the author (see appendix 3). On all occasions interviewers kept written records of each interview, thus allowing a check of apparently discrepant performances.

A further statistical check was made of inter-interviewer consistency during the Riverina testing, which preceded the Melbourne testing. An assumption was made that the population in government high schools in the Riverina was rather homogeneous and that there would not be significant differences between schools with respect to level of intellectual development. Accordingly, if different interviewers were assigned to each school and if the mean scores on the battery for each school showed significant differences, then this could indicate lack of consistency between interviewers. An analysis of variance

was performed on the scores for each school and there was no significant difference between them. Clearly this procedure is a dubious one because effect of interviewer is completely confounded with effect of school. It is reported, however, because the lack of significance lends support to the contention that inter-interviewer consistency was achieved for this study.

(ii) The Group Tests.

(a) Understanding in Science Test. This test was administered to all Ss by testers trained in its administration. The test is composed of four situations based on particular Piagetian tasks, for each of which is a series of questions. In their description of the administration of the test, Tisher and Dale (1975) suggest that, either Ss be shown the actual situation described in each part of the test before proceeding to the questions related to it, or that they be shown details of the situation by means of overhead transparencies. For this study Ss were shown the actual situation. At each administration of the test every S was allowed to complete a section before the groups proceeded to the next section. The tests were scored following the scheme described by Tisher and Dale (1975).

(b) The Sheehan Test. This test consists of 28 items but in the version employed in this study some wording changes were made on Sheehan's original version in order to minimise the American idiom. The language and complexity of the test makes it quite inappropriate for younger Ss and its use was restricted to Ss in years 10 to 12.

Scoring the Test. In the Understanding in Science test, items

which require the use of formal operations for their solution are given a score of two, items which only require the use of concrete operations for their solution are scored one. This procedure was adopted for scoring the Sheehan test, too, with one difference. In the Sheehan test there is a set of six items which require formal operations for their solution. They are in a multiple-choice format and the correct answer in each case involves choosing two of the possible five choices. For these questions, fully correct answers were scored two, answers which had chosen only one of the correct pair were scored one, and other responses were scored zero. The total score on the test was 44.

For the purpose of categorising the scores into the four broad Piagetian levels considered in this study, the following score ranges were arrived at following analysis of the test:

total score 0 - 12 = early concrete operations and below;
total score 13 - 22 = late concrete operations;
total score 23 - 35 = early formal operations;
total score 36 - 44 = complete formal operations.

A copy of the Sheehan test and scoring details may be obtained from the author.

2. Tests of Reading Comprehension. Examination of level of reading comprehension was restricted to the Riverina population. The instruments used were Co-operative Reading Comprehension Test Form L for subjects in years 7-10 (ACER, 1973) and the Co-operative Reading Comprehension Test Form Y for subjects in years 11 and 12 (ACER, 1975). Although these tests yielded sub-scale scores on speed, vocabulary and comprehension, for the purposes of this study only the compre-

hension scores have been used.

3. Test of Field Dependence-Independence (FDI). Examination of FDI was restricted to the Melbourne population. The criterion tests of this construct are individually-administered procedures, such as the embedded figures test (Witkin et al., 1971). Time and resources prevented the use of an individually-administered test, so the Group-Embedded Figures Test (GEFT) (Witkin et al., 1971) was used as a measure of FDI. This test has undergone preliminary validation, but is still subject to investigation.

The test was administered to all Ss in the Melbourne sample as a speed test, according to specifications described by the test's developers. Scoring followed the description provided by Witkin et al., (1971).

Estimating Validity of the Group Tests.

Cureton (1951) distinguished between three separate but related concepts concerning test validity.

(a) Cureton defined the predictive power of a test as the correlation between raw test scores and raw criterion scores.

(b) He defined the validity of a test as an estimate of the correlation between raw test scores and "true" (i.e. perfectly reliable) criterion scores.

(c) He defined the relevance of a test as an estimate of the correlation between "true" test scores and "true" criterion scores.

These three indices differ from one another because of the intrinsic unreliability of tests. Clearly, a test may have a high index of relevance but this may be only of theoretical

interest because its unreliability may reduce its predictive power.

In this study indices of predictive power have been reported, rather than indices of validity or relevance. The study was concerned to determine whether or not the two group tests employed are good predictors of level of intellectual development, using performance on Piagetian tasks as the criterion.

Determining the Significance of Correlation Coefficients

The usual procedures for determining the degree of significance of statistics, including the correlation coefficient, are based upon techniques developed for simple random samples. Ross (1976) has pointed out, however, that much social science research, especially survey research, is conducted with data obtained from the use of complex sample designs which depart from the assumption of simple random sampling. This study clearly is no exception.

The danger associated with the use of conventional procedures with data from complex sample designs lies in underestimation of sampling errors. Kish (1965) has defined a construct termed the "design effect" which can allow for the effect of the sample design in statistical procedures, and Ross (1976) has generated estimates of the "design effect" of various complex sampling designs for a range of statistics, including the correlation coefficient. This correction for design effect has a marginal effect on critical values for significance of correlation coefficients in this study, but it has been utilised where appropriate. In general, though, a study such as this which seeks strong relationships between

Table 5. Performance on individual Piagetian tasks and on total battery.

TASK	PERFORMANCE CATEGORY	PERCENTAGES	
		RIVERINA	MELBOURNE
Conservation of weight	≤IIa	13.1	12.9
	≥IIb	86.9	87.1
Conservation of displacement volume	≤IIa	23.0	22.0
	IIb	26.3	29.8
	≥IIIa	50.7	48.2
Separation of variables	IIa	11.2	11.5
	IIb	33.1	32.5
	IIIa	41.2	45.1
	IIIb	14.5	10.9
Equilibrium in the balance	IIa	8.8	8.5
	IIb	52.0	44.1
	IIIa	28.4	38.3
	IIIb	10.8	9.1
Total battery	IIa	12.9	10.5
	IIb	39.9	42.4
	IIIa	35.1	35.6
	IIIb	12.1	11.5

Table 6. Mean scores on the battery of Piagetian tasks for each school year in both populations.

School Year	Population					
	Riverina			Melbourne		
	Mean	Standard Deviation	No. in Sample	Mean	Standard Deviation	No. in Sample
7	11.8	2.1	54	12.2	2.0	52
8	12.7	1.7	54	12.8	1.9	64
9	12.6	2.1	59	13.2	1.9	52
10	13.7	1.8	48	13.4	2.0	63
11	14.3	1.9	53	13.9	1.9	41
12	14.9	1.5	28	14.3	1.7	23

measures as the bases for determining validity is not too concerned whether or not small values differ significantly from zero.

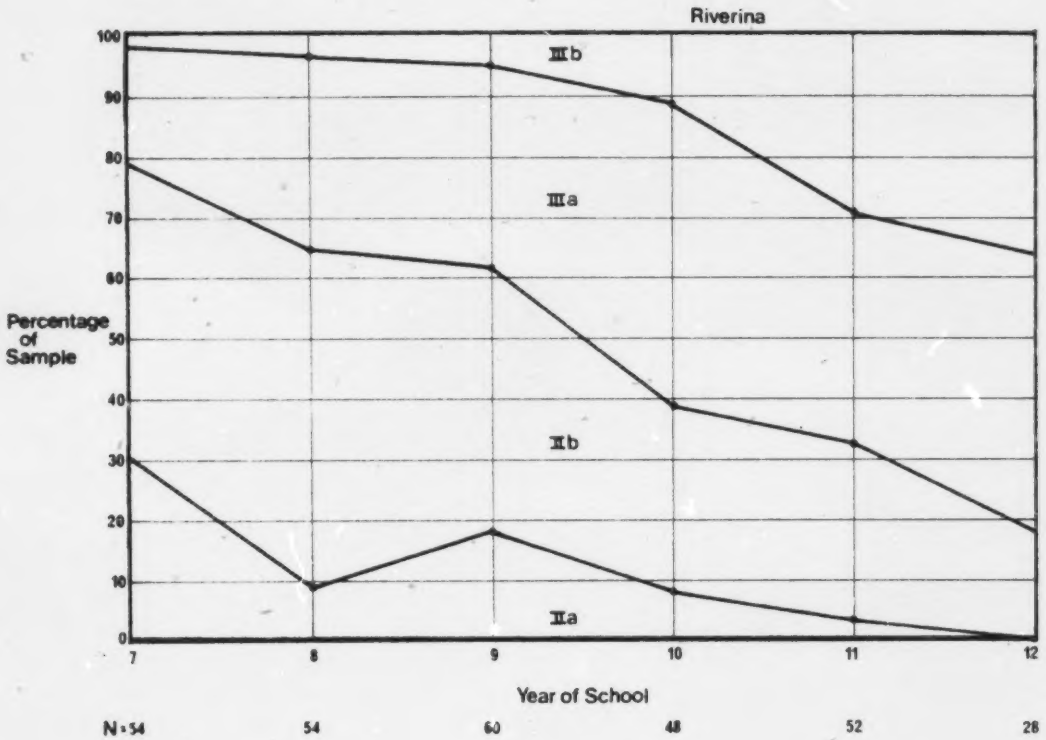
RESULTS AND DISCUSSION

Level of Intellectual Development in the Two Populations

The percentages of subjects operating at each Piagetian level on each of the four tasks, and on the total, are shown in Table 5. Table 6 presents the mean scores on the battery of tasks for each school year in both populations. The profiles across school years for both populations are presented in Figure 1.

A preliminary examination of the profiles in Figure 1 suggests that the populations differed in terms of the proportions of Ss at each Piagetian level at different school years. However, this is not the case. Inspection of mean scores for each year in Table 6 indicates no difference between populations for each school year and, in fact, all pairs of means were tested for significance and none was found. Furthermore, the percentages of Ss at each Piagetian level for each school year, for each population, are shown with 95% confidence limits in Table 7. Inspection of this table shows that the differences apparent in the profiles fall within these confidence limits and therefore must be regarded as not significant. Similarly, the performance on each task by Ss in the two populations is remarkably similar, as shown in Table 5.

These results indicate quite clearly that the majority of secondary school students in the two populations are not at the



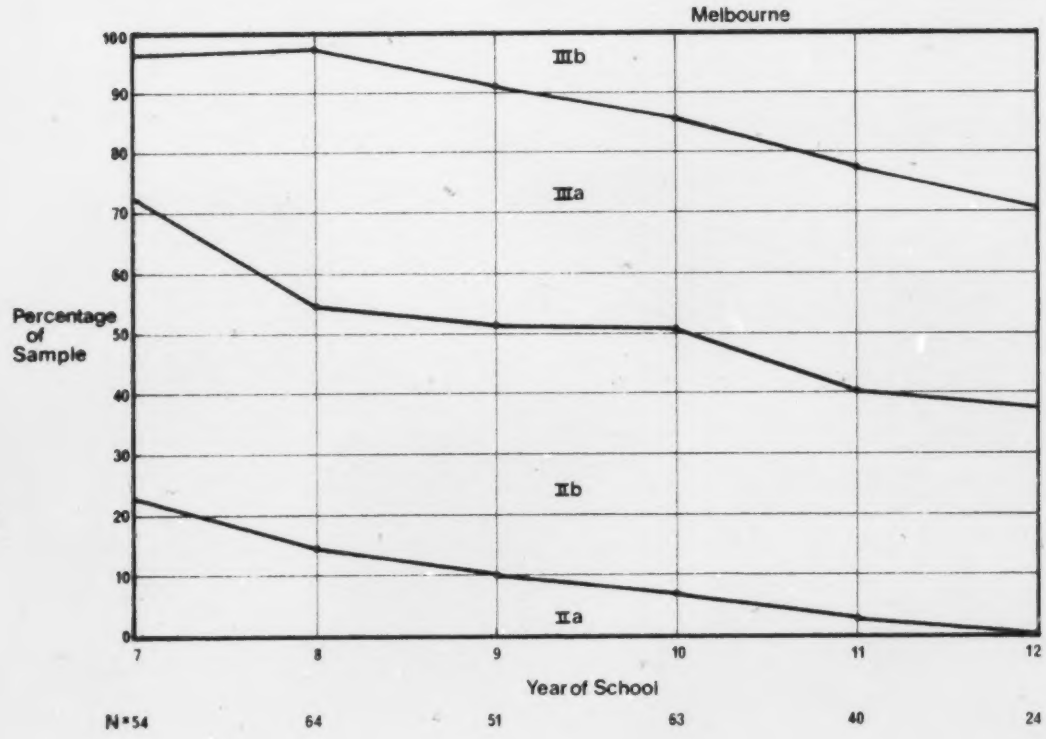


Figure 1 - Profiles of levels of intellectual development across school years for both populations.

Table 7. Percentages of subjects at each Piagetian level for each school year of both populations with 95% confidence limits.

Riverina Population

Piagetian Level	School Year					
	7	8	9	10	11	12
IIa	29.6 ± 12.2	9.3 ± 7.8	18.3 ± 9.8	8.3 ± 7.8	3.8 ± 5.2	0
IIb	50 ± 13.3	55.6 ± 13.3	43.3 ± 12.5	31.3 ± 13.1	28.8 ± 12.3	17.9 ± 14.2
IIIa	18.5 ± 10.4	31.5 ± 12.4	33.3 ± 11.9	50 ± 14.1	38.5 ± 13.2	46.4 ± 18.5
IIIb	1.9 ± 3.6	3.7 ± 5.0	5 ± 5.5	10.4 ± 8.6	28.9 ± 12.3	35.7 ± 17.8

Melbourne Population

Piagetian Level	School Year					
	7	8	9	10	11	12
IIa	22.2 ± 11.1	14.1 ± 8.5	9.8 ± 8.2	6.4 ± 6.0	2.5 ± 4.8	0
IIb	50 ± 13.3	40.7 ± 12.0	41.2 ± 13.5	44.4 ± 12.3	37.5 ± 15.0	37.5 ± 19.4
IIIa	24.1 ± 11.4	42.2 ± 12.1	39.2 ± 13.4	34.9 ± 11.8	35 ± 14.8	33.3 ± 18.9
IIIb	3.7 ± 5.3	2.1 ± 4.3	9.8 ± 8.2	14.3 ± 8.6	22.5 ± 12.9	29.2 ± 18.2

stage of formal operations, although by year 10 it may be stated that most show some predisposition towards formal thought. Such results are in accord with those obtained by other studies (e.g. Lovell, 1961; Dale, 1970; McKinnon and Renner, 1971; Higgens-Trenk and Gaite, 1971; Blake, 1976; Lawson and Blake, 1976) and indicate that the widely - held notion that formal operational thought develops in most humans at about eleven or twelve years of age is a myth.

It is interesting to note that if one considers the profiles in Figure 1 in terms of a concrete/formal dichotomy, the pattern presented is quite similar to that evolved by the ASEP developers (ASEP, 1974) on the basis of a little evidence and a lot of intuition. It is important to realise, though, that such a description of the population can be misleading. Although 20% - 30% of the Year 7 population may be classified loosely as "formal operational" ranging to above 70% - 80% being so classified at Year 12, (in accordance with the ASEP hypothetical profile) the great majority of Ss so classified only show some predisposition towards formal thought. That is, the majority of such Ss are at stage IIIa, according to this study at least, and cannot be regarded as completely formal operational thinkers. This is an important qualification and its implications are discussed later.

Relationship Between Sex and Performance on the Tasks

Most studies of performance on Piagetian tasks have shown that males perform at a superior level than females, or else that there is no difference between the sexes (e.g. Goldschmid, 1967; Goldschmid and Bentler, 1968; Friot, 1970). The relationship between performance on the tasks and sex in this study was analysed by calculating point biserial correlation

Table 8. Point biserial correlation coefficients between sex and performance on the Piagetian task battery.

Task	Point Biserial Correlations	
	Riverina	Melbourne
Conservation of weight	- .171*	- .049
Conservation of displacement volume	- .365*	- .261*
Separation of variables	- .325*	- .244*
Equilibrium in the balance	- .227*	- .220*
Total Score	- .388*	- .299*

* denotes significance at 0.01 level

coefficients for each task, and for the total score. These coefficients are shown in Table 8.

These correlations are rather low; one is not significant as shown in the table. However, they do indicate that males generally performed slightly but significantly better than females in both populations.

Field Dependence-Independence in the Melbourne Population

The mean scores for the group-embedded figures test (GEFT) for each school year, together with standard deviations and sample sizes, are presented in Table 9. Frequency distributions of GEFT scores for each school year and for the total sample are shown in Figure 2.

It is apparent that degree of field independence increases in the population quite markedly across the school years. Whereas almost 50% of Ss in year 7 exhibited, by their low GEFT scores, a significant level of field dependence, the majority of year 12 Ss showed a strong tendency towards field independence.

Relationship between FDI and Sex

Studies of FDI have shown that males generally are more field independent that females (e.g. Witkin, 1950; Witkin et al., 1954, 1962; Bieri et al., 1958; Gross, 1959; Goodenough, 1963). This was not the case in the Melbourne population, however. The point biserial correlation coefficient between sex and GEFT score was:

$r_{pb} = -0.097$, which is not significant.

The discrepancy evident between this result and those obtained in many other studies may be a consequence of the use

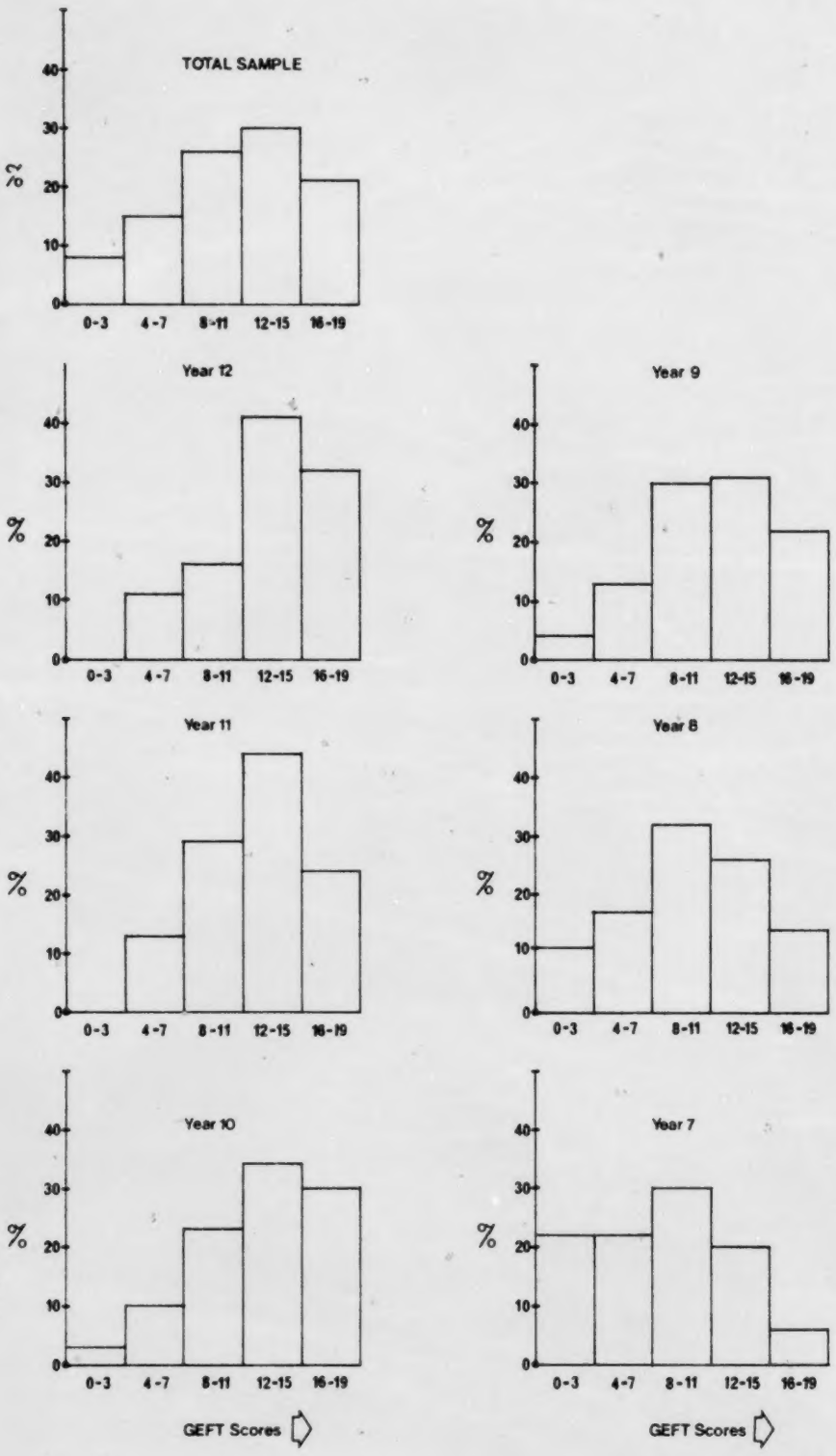


Figure 2 - Frequency distribution of GEFT scores across school years.

Table 9. Mean scores on the Group Embedded Figures Test across school years.

School Year	Mean Score	Standard Deviation	Sample Size
7	7.8	4.7	50
8	9.9	4.6	62
9	11.6	4.2	48
10	12.4	4.2	60
11	12.6	3.7	37
12	13.2	4.1	19

of the GEFT, rather than the individually-administered Embedded Figures Test, or another task which measures FDI.

Relationship Between FDI and Level of Intellectual Development

According to Pascual-Leone (1969), performance on Piagetian tasks is influenced by level of FDI, and the results indicate that both factors are affected by Ss age. In the context of this study, two questions emerged:

(a) is there a relationship between FDI and performance on the Piagetian tasks?

(b) if such a relationship exists, is it stronger than that which exists between performance on the tasks and age?

The correlation obtained between total score on the tasks and age was: $r = 0.323$. The correlation between total score on the tasks and score on the GEFT was: $r = 0.516$. Both of these coefficients are significant at the .01 level and, moreover, the difference between them is significant at the .001 level.

This means that, although performance on the Piagetian tasks is related to age of subject, it is more strongly related to level of FDI. This is in accord with Pascual-Leone's findings.

Validity of the Understanding in Science Test

An examination of the degree of relationship between performance on the battery of Piagetian tasks and performance on the Understanding in Science test was performed in two ways for each population. In the first test, an index of predictive

Table 10. Matrix of product-moment correlation coefficients between the Understanding in Science test and the Piagetian tasks for the Riverina Population.

Piagetian Tasks	Understanding in Science Test				
	Part I	Part II	Part III	Part IV	Total
conservation of weight	.245	.171	.202	.144	.298
conservation of displacement volume	.369	.235	.343	.151	.435
separation of variables	.407	.148	.339	.171	.429
equilibrium on the balance	.311	.266	.319	.173	.414
TOTAL	.467	.200	.425	.218	.549

* critical value (.05) = .134

* critical value established following the procedure of Ross (1976).

Table 11. Matrix of product-moment correlation coefficients between the Understanding in Science test and the Piagetian tasks for the Melbourne population.

Piagetian Tasks *	Understanding in Science Test				
	Part I	Part II	Part III	Part IV	Total
conservation of weight	.124(n.s.)	.153	.104	.005(n.s.)	.152
conservation of displacement volume	.281	.140	.304	.130(n.s.)	.322
separation of variables	.369	.265	.377	.106(n.s.)	.428
equilibrium in the balance	.333	.342	.230	.028(n.s.)	.382
TOTAL	.413	.326	.382	.107(n.s.)	.477

* critical value (.05) = .137

* critical value established following the procedure of Ross (1976).

power was calculated by obtaining the product-moment correlation coefficient between total score on the battery of tasks and total score on the Understanding in Science test. In addition to this coefficient, correlation coefficients were calculated between sections of the Understanding in Science test and the individual tasks.

The second test made no assumptions about the interval nature of the data. It was a chi-square test of independence between categorisations on the battery of Piagetian tasks and the Understanding in Science test. For each population a 4 x 4 contingency table was cast. The four categories derived from total scores on the Piagetian tasks were as defined in Table 4. The four equivalent categories for the Understanding in Science test scores were as given by Tisher and Dale (1975), viz.

0 - 10 early concrete,
11 - 19 late concrete,
20 - 26 early formal,
27 - 34 late (or complete) formal.

1. Indices of Predictive Power: Product-Moment Correlations

Matrices of correlation coefficients for both populations are presented in Tables 10 and 11. For both populations, the index of predictive power is relatively small. The difference between the two correlations - $r = 0.549$ for the Riverina and $r = 0.477$ for Melbourne - is not statistically significant. It is likely that corrections for attenuation for both the Understanding in Science test and the tasks would yield indices of validity and relevance considerably larger than this figure for predictive power. The fact remains, however, that users of the group test are concerned with its predictive power; and

Table 12. Chi-square values and contingency coefficients for tests of independence between categorisations on the Piagetian task battery and Understanding in Science test.

	Population	
	Riverina	Melbourne
χ^2	105 ($p < .0001$)	86 ($p < .0001$)
Contingency Coefficient	0.52	0.49

this is relatively low.

It is particularly illuminating to examine the relationship between part III of the Understanding in Science test and the equilibrium in the balance task (task 4). Part III of the Understanding in Science test is based on the equilibrium in the balance task. Consequently, if there was to be a strong relationship between the two tests, one would expect to see it between these components. The correlation, however, is not all strong - $r = 0.32$ for Riverina and $r = 0.23$ for Melbourne. These coefficients would be constrained by the restricted score ranges, particularly for the score on the task. However, it is still a clear indication that a written mode of test is not measuring the same thing that the corresponding Piagetian task is measuring.

2. Tests of Independence

In their initial assessment of the validity of the Understanding in Science test, Fisher and Dale (1975) performed a χ^2 test of independence. As one is concerned mainly with categorising Ss into Piagetian stages, rather than with actual raw scores, this seems to be a suitable means of testing the predictive power of the test.

The χ^2 values and corresponding contingency coefficients for both populations are presented in Table 12. These results illustrate a clear relationship between test and criterion, although it can be argued that it is not as strong as one would wish for a test which is purported to categorise Ss in the same way as the criterion does.

3. Factors Affecting the Relationship Between Test and Criterion.

(i) Level of Reading Comprehension

(a) One means of examining the extent to which level of reading comprehension may have influenced the relationship between performance on the Understanding in Science test and performance on the Piagetian tasks is as follows. If level of reading comprehension affects a S's ability to interpret the Understanding in Science test, then one may expect there to be a correlation between the size of the discrepancy between Understanding in Science scores and task battery scores as one variable, and level of reading comprehension as the other.

This analysis was performed by subtracting the total scores on the battery of tasks from the Understanding in Science total score for all Ss in the Riverina population and then calculating correlations with reading comprehension scores. Two correlations were calculated because of the two forms of the reading test.

They are:

Years 7 - 10 (Form L) $r = 0.44$ ($n = 201$; $p < .001$);

Years 11 - 12 (Form Y) $r = 0.24$ ($n = 79$; $p < .05$).

These results indeed indicate that discrepancy in performance between the Understanding in Science test and the criterion is related to level of reading comprehension. It would be anticipated that such a relationship would be stronger with younger Ss with lower levels of reading comprehension, as these results indicate.

(b) This relationship was subjected to further scrutiny by employing the contingency table used for the χ^2 test of independence between categorisations by tasks and by

Table 13. Means of reading comprehension scores for subjects in each cell of the contingency table of task categorisations against Understanding in Science Test categorisations.

Piagetian Task Categorisations	Reading Test Form	Understanding in Science Test Categorisations			
		IIa	IIb	IIIa	IIIb
IIa	L	11.93 (n=15)	16.3 (n=15)	- (n= 0)	- (n=0)
	Y	- (n= 1)	- (n= 1)	- (n= 0)	- (n=0)
IIb	L	14.78 (n=14)	15.05 (n=57)	20.00 (n=23)	- (n=0)
	Y	- (n= 1)	47.11 (n= 9)	52.11 (n= 9)	- (n=0)
IIIa	L	- (n= 0)	18.61 (n=28)	21.10 (n=31)	- (n=0)
	Y	- (n= 0)	50.82 (n=11)	55.35 (n=23)	- (n=0)
IIIb	L	- (n= 0)	15.25 (n= 4)	24.60 (n= 5)	- (n=1)
	Y	- (n= 0)	51.50 (n= 4)	56.00 (n=15)	59.00 (n=4)

(Numbers in parentheses are sample sizes)

the Understanding in Science test. Reading comprehension scores were extracted for each S in each cell of this contingency table and, where possible, means for reading comprehension were calculated. This was not feasible for some cells of the contingency table because there was either no subjects or only one subject in the cell. The table containing the means is presented in Table 13.

Examination of these means suggests that Ss categorised at the same Piagetian stage by the tasks, but differentiated into different stages by the Understanding in Science test, exhibit concomitant differences in reading comprehension. It appears, in fact, as though superior reading comprehension results in superior performance on the Understanding in Science test; Piagetian task performance being equal.

Tests of significance of the differences between means within one task category (i.e. across the rows of the table) were performed where sample size allowed. These tests were performed as follows:

(1) For a test involving only two means, a one-tailed t test was performed to test these hypothesis that $\mu_1 > \mu_2$;

(2) for a test involving more than two means, an analysis of variance was performed, followed where appropriate by a Newman-Keuls test for differences between all pairs of means.

The results of these tests are shown in Table 14. A distinct trend is evident in these results, supporting the contention that the relationship between performance on the Understanding in Science test and performance on a criterion battery of Piagetian tasks is influenced by level of reading comprehension.

Table 14. Results of tests of significance of differences between means of reading comprehension scores within rows of the tasks categorisation, Understanding in Science Test categorisation contingency table.

Piagetian Task Categorisations	Reading Test Form	Understanding in Science Test Categorisations			
		IIa	IIb	IIIa	IIIb
IIa	L	11.93—*——16.13			
IIb	L	14.78—n.s.—15.05—*—20.00			
	Y		47.11—*—52.11		
IIIa	L		18.61—*—21.10		
	Y		50.82—*—55.35		
IIIb	L		15.25—*—24.60		

* denotes significant difference at 0.05 level (1-way test)

n.s. denotes not significant at 0.05 level

(ii) Field Dependence-Independence

For the Melbourne population which was tested with the GEFT, the index of predictability for the Understanding in Science test was 0.47. The possible influence of FDI on this correlation was ascertained by partialling out the effect due to FDI. That is, a partial correlation coefficient was calculated, as shown below.

$r_{TU} = 0.47$; $r_{TU.G} = 0.26$

where T = score on tasks,

U = score on Understanding in Science,

G = GEFT score.

Partialling out the correlation due to the effect of FDI decreased the correlation between the group test and its criterion from 0.47 to 0.26. This means that approximately 45% of the association between performance on the tasks and performance on the Understanding in Science test results from the effect of FDI.

Validity of the Sheehan Test

As for the Understanding in Science test, an examination of the relationship between performance on the Sheehan test and performance on the battery of Piagetian tasks was made in two ways. Firstly, indices of predictive power were calculated from the raw scores and secondly, χ^2 tests of independence were performed for both populations.

1. Indices of Predictive Power: Product - Moment Correlations.

The indices of predictive power that were calculated for each population are as follows:

Table 15. Chi-square values and contingency coefficients for tests of independence between categorisations on the Piagetian task battery and on the Sheehan test.

	Population	
	Riverina	Melbourne
χ^2	27.3 ($p < .01$)	17.7 ($p < .05$)
Contingency Coefficient	0.43	0.37

Riverina - $r = 0.39$ ($n = 121$, $p < .001$);

Melbourne - $r = 0.38$ ($n = 113$, $p < .001$).

There obviously is a relationship between the test and the criterion, but the predictive power of the test is quite low. It is stronger, however, than that observed by Blake (1976) when the test was used with first-year college students.

2. Tests of Independence. The raw scores on both Sheehan test and total tasks score were cast in a 4 x 4 contingency table, employing categorisations previously described. The χ^2 values and corresponding contingency coefficients are presented in Table 15. Again, evidence of a relationship between the two tests is evident, but it is not as strong as one would expect for a highly predictive test.

3. Factors Affecting the Relationship Between Test and Criterion.

(i) Level of Reading Comprehension.

(a) As was the case with the Understanding in Science test, one way of investigating the relationship between test, criterion and level of reading comprehension is to examine relationships between level of reading comprehension and the discrepancy between scores on the tasks and scores on the Sheehan test.

This analysis was performed by subtracting the total scores on the battery of Piagetian tasks from the Sheehan test scores and then calculating correlations with reading comprehension scores. The correlations for the two forms of the reading test were:

Year 10 (Form L) $r = 0.61$ ($n = 41$; $p < .01$);

Years 11 - 12 (Form Y) $r = 0.35$ ($n = 77$; $p < .01$).

As the Sheehan test was only used with years 10 - 12, the

sample sizes are rather small. However, these results suggest quite a strong relationship between discrepancy of performance on the two tests of level of intellectual development and level of reading comprehension. This effect appears to be stronger than with the Understanding in Science test and this is to be expected, given that the Sheehan test requires a great deal of reading and comprehension in comparison with the Understanding in Science test.

(b) The relationship was examined further by investigating the reading comprehension scores of Ss in each cell of the 4 x 4 contingency table of categorisations based on the Sheehan test against categorisations based on the Piagetian tasks. The procedure followed was the same as that used for the similar analysis of the Understanding in Science test. Table 16 presents the mean reading comprehension scores for the Riverina population. It is evident that the numbers in most cells are so small as to make any overall analysis of trends out of the question. However, four tests of significance between means within rows were performed in the manner described earlier. These results are shown in Table 17. It is clear that the trend that is evident supports the contention that Ss equally categorised by the battery of tasks, but differentiated by the Sheehan test, show a trend in level of reading comprehension. Unfortunately, the small numbers have rendered the testing impractical, and the non-significant results shown in Table 17 could be a consequence of these small sample sizes.

(ii) Field Dependence-Independence. For the Melbourne sample, the index of predictability of the Sheehan test was 0.38. When the effect of GEFT score is partialled out, the

Table 16. Means of reading comprehension scores for subjects in each cell of the contingency table of task categorisations against Sheehan test categorisations.

Piagetian Task Categorisations	Reading Test Form	Reasoning Test Categorisations IIa	IIb	IIIa	IIIb
IIa	L	- (n=1)	- (n=1)	- (n= 1)	- (n = 0)
	Y	- (n=1)	- (n=0)	- (n= 1)	- (n= 0)
IIb	L	- (n=0)	17.67 (n=6)	21.50 (n= 4)	- (n= 1)
	Y	- (n=1)	- (n=1)	51.76 (n=16)	55.00 (n= 2)
IIIa	L	- (n=1)	16.25 (n=4)	22.00 (n=11)	23.50 (n= 2)
	Y	- (n=1)	48.00 (n=2)	52.67 (n=18)	55.58 (n=12)
IIIb	L	- (n=0)	- (n=1)	23.25 (n= 4)	- (n= 0)
	Y	- (n=0)	- (n=1)	55.27 (n=11)	56.73 (n=11)

(Numbers in parentheses are sample sizes)

Table 17. Results of tests of significance of differences between means of reading comprehension scores within rows of the tasks categorisation, Sheehan test categorisation contingency table.

Piagetian Test Categorisations	Reading Test Form	Reasoning Test Categorisations			
		IIa	IIb	IIIa	IIIb
IIa	L				
	Y				
IIb	L		17.67——*—21.50		
	Y				
IIIa	L		16.25--n.s.—22.00—n.s.—23.50		
	Y		48.00—n.s.—52.67—n.s.—55.58		
IIIb	L				
	Y			55.27—n.s.—56.73	

* denotes significant difference at 0.05 level (1-way test)

n.s. denotes not significant at 0.05 level.

result is as follows:

$r_{TS} = 0.38$; $r_{TS.G} = 0.15$

where T = tasks score,

S = Sheehan test score,

G = GEFT score.

Hence, partialling out the effect due to FDI drops the correlation between tasks score and Sheehan test score quite considerbaly. This means that approximately 61% of the association between performance on the tasks and performance on the Sheehan test results from the effect of FDI.

CONCLUSIONS AND IMPLICATIONS

Level of Intellectual Development in the Australian Secondary School Population

The fact that the profiles for the two populations, representing different state education systems and different social settings, are so alike suggests that the pattern of intellectual development evident in the profiles may be generalizable to the wider Australian context. Arons and Karplus (1976), in reviewing studies of intellectual development, struck a pessimistic note regarding the success of American secondary education in facilitating growth towards formal operational thought. This study indicates that growth does occur in Australian secondary schools, although it is important to note that not all of the change evident in the profiles in Figure 1 may be due to growth. It is quite conceivable that some of the change, especially after Years 9 and 10, may be due to the differential drop-out of concrete operational students from school. It is intended to investi-

gate further this possibility by means of a longitudinal study.

The major purpose of education, in New South Wales at least, has been made quite clear. It is to promote the development of individuals in all its aspects; intellectual, moral, affective. The implicit outcome of such a goal is in accord with the statement made by the United States Educational Policies Commission in 1961 (Educational Policies Commission, 1961); that education should foster the development of rational thought. In Piagetian terms, this means to foster the development of formal operational thought. Such being the case, the results of this study have particular implications for the secondary curriculum.

Firstly, teachers and curriculum developers need to be aware that a significant proportion of the secondary school population is composed of concrete operational thinkers who cannot handle abstract reasoning. That is, many students cannot handle hypothetical arguments and propositions which are not rooted in concrete reality, including the hypothetico-deductive procedures of science, and they cannot cope effectively with combinatorial logic. Moreover, the great majority are only beginning to show early competence with formal thinking. It is important, therefore, for teachers and curriculum developers to analyse their materials and procedures in order to establish that they are not making unwarranted assumptions about the logical capabilities of their students. In essence, this is what ASEP attempted in constructing its materials at three levels.

The implications of these results extend beyond making teachers and curriculum developers aware that most adolescents are not formal thinkers, however. If the seminal work of

Inhelder and Piaget (1958) indicated that a high proportion of Genevan adolescents were formal operational, at least by about sixteen years of age, why have these findings not been replicated in other studies, including this one? Piaget (1972) conceded that the original Genevan sample may have been highly selective and certainly the more recent study of the British school population by Shayer et al. (1976) indicates that academically-selected subjects appear intellectually more advanced on Piagetian tests than their unselected peers. The point is, however, that it is possible that such apparently gifted students may represent a rate of intellectual development which is optimal for most, and that the general rate of growth, as evidenced in the results of this study, could result from retardation.

Inhelder and Piaget (1958) pointed out that "a particular social environment remains indispensable for the realization of these possibilities. It follows that their realization can be accelerated or retarded as a function of cultural and educational conditions" (p.237). It is conceivable that our secondary schools often do not provide educational conditions appropriate for realizing the intellectual growth of their students. The desirable conditions for such growth involve more than teachers being aware of the existing cognitive capacities of students, as discussed above. They involve the use of materials and strategies which set out to promote intellectual growth. In the area of science education, the implications of such an approach have been described by Renner and Lawson (Renner and Lawson 1973a, 1973b; Lawson and Renner, 1975; Renner, 1976). More thought needs to be given to the problem at a general curriculum level. If formal operational thought is achievable

by most adolescents by the time they reach sixteen to eighteen years of age, given the appropriate cultural and educational settings, then the results of this study suggest that our schools are not providing such conditions.

The Applicability of the Understanding in Science Test and the Sheehan Test

As a measure of stage or level of intellectual development, the Understanding in Science test does not have very high predictive power. Consequently its use in place of individually-administered Piagetian tasks is not recommended.

The Sheehan test is even less suitable as a measure of level of intellectual development. It appears to be as much or more a measure of field dependence-independence as it is of level of intellectual development.

The study makes clear some of the difficulties in producing a group test designed to measure what the Piagetian tasks measure. The change of mode from oral exchange and manipulation of apparatus to written mode creates at least two difficulties. It introduces the factor of reading comprehension which can differentially affect performance on such a group test. It also alters the context in which the subject is confronted with the problem and this seems to cause an alteration in the way in which the subject's level of field independence affects performance on the test.

Finally, it should be stressed that the Understanding in Science test and the Sheehan test could be useful tests in their own rights. Indeed, they may be better predictors of success at school than performance on Piagetian tasks. It is

that they are not very good predictors of performance on Piagetian tasks and, consequently, of stage of intellectual development.

REFERENCES

Arons, A.B. and Karplus, R. Implications of accumulating data on levels of intellectual development. American Journal of Physics, 1976, 44(4), 396.

Australian Council for Education Research. Co-operative Reading Comprehension Test: Form Y. Hawthorn, Victoria: ACER, 1975.

Australian Council for Educational Research. Co-operative Reading Comprehension Test: Form L. Hawthorn, Victoria: ACER, 1973.

Australian Science Education Project. A Guide to ASEP. Melbourne: State Govt. of Victoria, 1974.

Bieri, J., Bradburn, W.M. and Galinsky, M.D. Sex differences in perceptual behaviour. Journal of Personality, 1958, 26, 1-12.

Blake, A.J.D. Relationships Between Science Instruction, Cognitive Style and the Development of Formal Operations in Australian College Students. Unpublished doctoral thesis, Purdue University, 1976.

Blake, A.J.D., Lawson, A.E. and Nordland, F.H. The Karplus islands puzzle: Does it measure Piagetian operations? Journal of Research in Science Teaching, 1976, 13(5), 397-404.

Case, R. and Globerson, T. Field-independence and central computing space. Child Development, 1974, 45, 772-790.

Cureton, E.E. Validity. In, Lindquist, E.F. (ed.), Educational Measurement. Washington: American Council on Education, 1951.

Dale, L.G. The growth of systematic thinking: Replication and analysis of Piaget's first chemical experiment. Australian Journal of Psychology, 1970, 22(3), 277-286.

Educational Policies Commission. The Central Purpose of American Education. Washington: National Education Association, 1961.

Friot, F.E. The Relationship Between an Inquiry Teaching Approach and Intellectual Development. Unpublished doctorial dissertation, University of Oklahoma, 1970.

Goldschmid, M.L. Different types of conservation and nonconservation and their relation to age, sex, IQ, MA and vocabulary. Child Development, 1967, 38, 1229-1246.

Goldschmid, M.L. and Bentler, P.M. Manual: Concept Assessment Kit - Conservation. San Diego: Educational and Industrial Testing Service, 1968.

Goodenough, D.R. Sex differences in psychological differentiation. Paper presented at the biennial meeting, Society for Research in Child Development, 1963.

Gross, F. The role of sex in perception of the upright. Journal of Personality, 1959, 27, 95-103.

Higgens-Trenk, A. and Gaite, A.J.H. Elusiveness of formal-operational thought in adolescents. Proceedings 79th Annual Convention APA, 1971, 201-202.

Ingle, R.B. and Shayer, M. Conceptual demands in Nuffield O-level chemistry. Education in Chemistry; 1971,8, 182-183.

Inhelder, B. and Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Karplus, E. and Karplus, R. Intellectual development beyond elementary school I: Deductive logic. School Science and Mathematics, 1970, 70(5), 398-406.

Karplus, R. and Lavatelli, C. The Developmental Theory of Jean Piaget: Conservation. (Film). San Francisco: John Davidson Film Producers, 1969.

Karplus, R. and Peterson, R.W. Intellectual development beyond elementary school II: Ratio, a survey. School Science and Mathematics, 1970, 70(6), 813-820.

Karplus, R. and Thier, H.D. A New Look at Elementary School Science. Chicago: Rand-McNally, 1967.

Keeves, J.P. and Bourke, S.F. Australian Studies in School Performance. Volume 1, Literacy and Numeracy in Australian Schools: A First Report. Canberra: ERDC, 1976.

Kish, L. Survey Sampling. New York: John Wiley, 1965.

Lawson, A.E. Relationships Between Concrete and Formal Operational Science Subject Matter and the Intellectual Level of the Learner. Unpublished doctoral dissertation, University of Oklahoma, 1973.

Lawson, A.E. and Blake, A.J.D. Concrete and formal thinking abilities in high school biology students as measured by three separate instruments. Journal of Research in Science Teaching, 1976, 13(3), 227-235.

Lawson, A.E. Blake, A.J.D. and Nordland, F.H. Training effects and generalization of the ability to control variables in high school biology students. Science Education, 1975, 59(3), 387-396.

Lawson, A.E. and Renner, J.W. Piagetian theory and biology teaching. American Biology Teacher, 1975, 37(6), 336-343.

Longeot, F. Un essai d'application de la psychologie genetique a la psychologie differentielle. Bulletin de l'institut National D'Etude, 1962, 18, 153-162.

Longeot, F. Analyse statistique de trois tests genetiques collectifs. Bulletin de l'institut National D'Etude, 1965, 20, 219-237.

Lovell, K.A. A follow-up study of Inhelder and Piaget's "The Growth of Logical Thinking". British Journal of Psychology, 1961, 52, 143-153.

McKinnon, J.W. and Renner, J.W. Are colleges concerned with intellectual development? American Journal of Physics, 1971, 39, 1047-1052.

Niemark, E.D. Longitudinal development of formal operations. Genetic Psychology Monographs, 1975, 91, 171-225.

Pascual-Leone, J. Cognitive Development and Cognitive Style: A General Psychological Integration. Unpublished doctoral thesis, University of Geneva, 1969.

Piaget, J. Intellectual evolution from adolescence to adulthood. Journal of Human Development, 1972, 15, 1-12.

Piaget, J. and Inhelder, B. Le Developpement des Quantites Physique chez l'Enfant. Neuchatel: Delachaux et Niestle, 1941.

Renner, J.W. Significant physics content and intellectual development - cognitive development as a result of interacting with physics content. American Journal of Physics, 1976, 44(3), 218-222.

Renner, J.W. and Lawson, A.E. Piagetian theory and instruction in physics. The Physics Teacher, 1973a, 11(3), 165-169.

Renner, J.W. and Lawson, A.E. Promoting intellectual development through science teaching. The Physics Teacher, 1973b, 11(5), 273-276.

Rosier, M.J. and Williams, W.H. The Sampling and Administration of the IEA Science Project in Australia 1970: A Technical Report. IEA (Australia) Report 8. Hawthorn, Victoria: ACER, 1973.

Ross, K.N. Searching for Uncertainty: Occasional Paper No. 9 Hawthorn, Victoria: ACER, 1976.

Rowell, J.A. and Hoffmann, P.J. Group tests for distinguishing formal from concrete thinkers. Journal of Research in Science Teaching, 1975, 12(2), 165-175.

Schools Council 5/13 Project. With Objectives in Mind. London: Macdonald Educational, 1972.

Shayer, M. How to assess science courses. Education in Chemistry, 1970, 7, 182-186.

Shayer, M. Conceptual demands in Nuffield O-level physics course. School Science Review, 1972, 54, 26-34.

Shayer, M. and Wharry, D. Piaget in the classroom part I: Testing a whole class at the same time. School Science Review, 1974, 55, 447-457.

Shayer, M., Küchemann and Wylam, H. The distribution of Piagetian stages of thinking in British middle and secondary school children. British Journal of Educational Psychology, 1976, 46, 164-173.

Sheehan, D.J. The Effectiveness of Concrete and Formal Instructional Procedures with Concrete - and Formal - Operational Students. Unpublished doctoral dissertation, State Univ. of New York at Albany, 1970.

Tisher, R.P. and Dale, L.G. Understanding in Science Test. Hawthorn, Victoria: ACER, 1975.

Witkin, H.A. Individual differences in ease of perception of hidden figures. Journal of Personality, 1950, 19, 1-15.

Witkin, H.A., Dyk, R.B., Faterson, H.F., Goodenough, D.R. and Karp, S.A. Psychological Differentiation: Studies of Development. New York: Wiley, 1962.

Witkin, H.A., Lewis, H.B., Hertzman, M., Machover, K., Meissner, P.B. and Wapner, S. Personality through Perception: An Experimental and Clinical Study. New York: Harper, 1954.

Witkin, H.A., Oltman, P.K., Raskin, E. and Karp, S.A. A Manual for the Embedded Figures Tests. Palo Alto: Consulting Psycholgists Press, 1971.

APPENDICES

APPENDIX 1

Schools included in the Riverina sample and details of sample structure

School	Sample Structure Year and Sex												TOTAL	
	7		8		9		10		11		12			
	M	F	M	F	M	F	M	F	M	F	M	F	M	F
Albury H.S.	2	1	2	1	2	2	2	2	2	2	2	1	12	9
Albury North H.S.	2	2	2	2	2	2	2	0	2	3	2	2	12	11
Coleambally C.S.	1	2	1	1	2	1	0	0	0	0	0	0	4	4
Cootamundra H.S.	2	2	2	2	2	1	1	1	1	1	0	0	8	7
Corowa H.S.	1	0	1	1	1	1	1	1	2	2	1	2	7	7
Finley H.S.	2	2	2	2	2	1	1	2	3	2	0	0	10	9
Griffith H.S.	1	1	1	1	2	2	2	2	1	1	1	1	8	8
Junee H.S.	0	0	0	0	1	1	2	2	2	2	1	1	6	6
Kooringal H.S.	2	2	2	2	2	2	2	1	0	0	0	0	8	7
Leeton H.S.	1	2	2	2	2	2	2	2	2	0	2	1	11	9
Mount Austin H.S.	2	2	2	2	1	1	1	1	2	2	1	2	9	10
Mount Erin H.S.*	-	3	-	3	-	2	-	2	-	-	-	-	-	10
Murrumburrah H.S.	2	1	2	1	1	1	0	0	0	0	0	0	5	3
Narrandera H.S.	1	1	1	1	2	1	2	1	2	1	2	0	10	5
St. Michael's H.S.*	3	-	3	-	2	-	2	-	-	-	-	-	10	-
St. Paul's College*	1	1	1	1	1	1	1	1	1	1	0	0	5	5
Temora H.S.	2	0	1	1	1	1	1	1	1	1	1	1	7	5
Trinity Senior H.S*	-	-	-	-	-	-	-	-	5	5	0	0	5	5
Tumut H.S.	1	1	2	1	2	2	1	2	0	2	1	1	7	9
Wade H.S.	1	1	1	1	2	2	2	2	1	1	1	1	8	8
Wagga Wagga H.S.	2	2	2	2	2	2	2	1	2	1	1	2	11	10

* denotes non-government schools

APPENDIX 2

Schools included in the Melbourne sample and details of sample structure

School	Sample Structure Year and Sex												TOTAL	
	7		8		9		10		11		12			
	M	F	M	F	M	F	M	F	M	F	M	F	M	F
Bentleigh H.S.	1	4	2	2	0	3	0	4	2	1	0	1	5	15
Broadmeadows West T.S.	4	-	5	-	2	-	5	-	4	-	-	-	20	-
Chavoin College*	-	5	-	5	-	6	-	1	-	1	-	2	-	20
Era School*	3	0	2	2	3	2	2	0	0	2	3	2	13	8
Fawkner H.S.	2	2	3	1	1	2	1	2	1	2	1	2	9	11
Glen Waverley H.S.	1	2	1	3	2	0	3	1	2	1	2	2	11	9
Kingsbury T.S.	5	-	5	-	2	-	5	-	3	-	-	-	20	-
Montmorency H.S.	1	1	2	3	4	1	2	3	1	1	0	1	10	10
Pembroke H.S.	1	2	2	4	0	3	2	0	1	2	1	2	7	13
Reservoir H.S.	1	2	2	3	1	1	2	3	2	1	2	0	10	10
St.Columb a's College*	-	4	-	1	-	3	-	5	-	6	-	1	-	20
Sunshine West H.S.	2	0	1	2	2	0	2	2	1	2	0	2	8	8
Templestowe T.S.	1	0	1	3	5	2	2	5	1	0	-	-	10	10
Therry College*	5	-	5	-	5	-	5	-	0	-	0	-	20	-
Whitehorse Tech. College*	-	3	-	4	-	3	-	6	-	4	-	0	-	20

* denotes non-government schools

APPENDIX 3

List of trained interviewers employed in administration of the Piagetian tasks.

W. Adams

J. Allen

J. Baird

A. Blake

P. Bryant

R. Doon

J. Fife

D. Finch

D. Gagg

M. Gooch

R. Hamilton

M. Healy

D. Hill

N. Inglis

P. McPhee

M. Muller

G. O'Sullivan

W. Palframan

A. Roben

R. Schippers

S. Thick